Aligning biogHist in both EAS

The element <biogHist>, used to record biographical or historical information - mainly - about the creator(s) of archival materials, has always been one of the main crossovers between EAC-CPF describing those agents themselves in more detail and EAD describing the materials they created, maintained and used. It is hence, next to the <control> section, the biggest aspect of reconciliation between the two EAS as they currently stand.

Current status

The status summarised below represents what is currently published via the EAC-CPF and EAD3 schemas and tag libraries as well as taking into account those decisions that have been taken already during EAC-CPF team as well as some joint team meetings between EAC-CPF, EAD and Schema teams during the last few months.

Spelling

One of the differences between the EAS that is more directly visible is the spelling, with EAC-CPF using camelCase and EAD3 continuing to use all lowercase spellings. After initial conversations during the annual meeting in Austin, Texas, in 2019 as well as following conversations during the Berlin meeting of the EAC team in March 2020, Schema team has led two rounds of survey with all TS-EAS members to decide the question of spelling:

- The first round included six options, which - next to the existing ones for EAC-CPF and EAD3 - also covered kebab-case, snake_case, PascalCase, and UPPERCASE. TS-EAS members were asked to rank these six options according to their preference from most to least preferred. While UPPERCASE was the least preferred option by large, followed by PascalCase and snake_case, there was no clear preference between the remaining three options with the results of this first survey.
- Schema team therefore decided to have a second survey, which asked for a definitive vote for one of the following three: camelCase, lowercase, or kebab-case. While there still was a tie between camelCase and kebab-case, the most significant result of this second vote was the fact, that out of 13 votes none was cast for lowercase. Given that the EAS family already includes an example for using camelCase with EAC-CPF, which also had a minimal lead over the option of kebab-case, the decision was taken to move forward with using camelCase for all EAS in future.

Relevant decision in the context of the current document

The spelling convention of camelCasing will be used for all EAS in future, i.e. the **tag name will be <biogHist>**.

Attributes

In terms of attributes used with <biogHist>, the current differences between EAC-CPF and EAD3 will be solved mostly by the general changes around attributes for both EAS that already have been discussed as part of the EAC-CPF revision.

- EAC-CPF will remove the XML namespace and @xml:id will be used as @id as already is the case in EAD3.
- EAC-CPF will introduce the attribute @audience with all elements as part of EAD3 reconciliation.
- EAC-CPF will use language attribution via the attributes @languageOfElement and @scriptOfElement on all non-empty elements, i.e. all elements that either can have text themselves or via their sub-elements.
    - EAD3 will ensure that this same approach is followed through, including the renaming of the current attributes @lang and @script.
- **Pending:** <biogHist> in both EAS currently uses the attribute @localType. While this means that there is no need for reconciliation, there is a general question pending with regard to the use of @localType (to be discussed on 6 August during the TS-EAS Annual Meeting).
- **Pending:** <biogHist> in EAD3 currently also uses the attributes @altrender and @encodinganalog. Especially @altrender, used for formatting purposes, but also @encodinganalog, used to provide the numbers/codes/identifiers of equivalent elements or fields in related standards such as ISAD(G), will be reviewed as part of the upcoming major revision of EAD.

Relevant decision in the context of the current document

<biogHist> will use the **attributes @audience, @id, @languageOfElement, and @scriptOfElement** as well as, potentially, @localType.

During the meeting on 6 August, TS-EAS decided to keep @localType and to mainly review the use of the attribute within <control>. **<biogHist>** will hence **continue using @localType**.

Sub-elements

The main differences remain with regard to the sub-elements available with <biogHist> in both EAS at the moment as well as the sub-elements of those elements themselves. There are three groups of sub-elements to be considered in the context of this document:

- Elements that are available in both EAS;
- Elements that are used only in one EAS, though generally would be available in the other;
- Elements that are used only in one EAS and are not available in the other.

Elements available in both EAS

<chronList>

Contains <chronItem> in both EAS, but also contains <head> and <listhead> in EAD3. The sub-element <chronItem> itself includes date elements, <event> and an option to encode a place in both EAS, with EAD3 additionally using <chronItemSet> to bundle one date with several events and/or places. Furthermore, EAD3 allows for all three date elements, while EAC-CPF only has <date> and <dateRange>.

Suggestion

Allow for all three date elements to be used with <chronItem> in EAC-CPF and introduce **<chronItemSet> as a new sub-element**. Find technical solution in the schema definition for additional formatting and structuring elements, such as <list> and <listhead>.

Decision

During the meeting on 6 August, TS-EAS confirmed the **addition of <dateSet> and <chronItemSet>** as sub-elements of <chronItem> in EAC-CPF.

<list>

Only has <item> as sub-element in EAC-CPF, while it also contains <defitem>, <head>, and <listhead> in EAD3. Furthermore, the sub-element <item> in EAD3 comes with an extended mixed content model.

Suggestion

Deal with mixed content difference via schema definition and consider the option of <span@localType> instead of a more complex content model when revising EAD.

Decision

During the meeting on 6 August, TS-EAS confirmed that mixed content difference will be **dealt with technically** via schema definition.

<p>

Uses an extended mixed content model in EAD3, but only in EAC-CPF.

Suggestion

Deal with mixed content difference via schema definition and consider the option of <span@localType> instead of a more complex content model when revising EAD.

Decision

During the meeting on 6 August, TS-EAS confirmed that mixed content difference will be **dealt with technically** via schema definition.

Elements used in one EAS only, though available in the other

<citation>

Is only used in EAC-CPF as a sub-element of <biogHist>, while EAD3 uses <ref> as sub-element of <biogHist><p>.

Decision - part 1

During the TS-EAS Annual Meeting on 4 August it was decided to replace <citation> with <ref> and to rename the element to <reference>.

Suggestion

Use <reference> as sub-element of <p> rather than directly as sub-element of <biogHist>.

Decision - part 2

During the meeting on 6 August, TS-EAS confirmed that **<reference> will be used within <p>** rather than directly within <biogHist> in both EAS.

<abstract>

Only used with <biogHist> in EAC-CPF, while EAD3 uses <abstract> only as sub-element of <did>.

Suggestion

Review the potential use of <abstract> as a dedicated summary with <biogHist> and other informal descriptive elements such as <scopeContent> as part of the major revision of EAD.

Decision

During the meeting on 6 August, TS-EAS confirmed that this will be **looked at during the major revision of EAD**.

<biogHist>

Same as the other informal descriptive elements, <biogHist> can be nested within itself in EAD.

Suggestion

Understand use cases and need for such nesting with the potential end to either remove it or consider alternative options. E.g. allowing <head> to be repeatable if this is mainly about having different levels of headers in a longer descriptive text.

Decision

During the meeting on 6 August, TS-EAS confirmed that this will be **looked at during the major revision of EAD**.

Elements used in one EAS

Additional formatting elements <blockquote>, <head>, and <table>

EAD3 - in general - allows for more formatting and structuring than EAC-CPF currently does and is likely to require. In the context of <biogHist> this refers to the sub-elements of <blockquote>, <head>, and <table>.

Suggestion

Deal with this difference via schema definition.

Decision

During the meeting on 6 August, TS-EAS confirmed that mixed content difference will be **dealt with technically** via schema definition.

<outline>

EAC-CPF includes <outline> for <biogHist>, which does not exist in EAD3.

Suggestion

Review use cases and usefulness of the element <outline> in general with the potential end of removing it from EAC-CPF.

Decision

During the meeting on 6 August, TS-EAS confirmed that this will be **reviewed at one of the next meetings of the EAC team**. Mark will have a closer look at <outline> examples drawn from the sample data sets he has analysed.

Follow up:

Two files:

- eac:outline.xml: which contains the 6 extracted encodings using “outline”. 5 are from SNAC, and 1 is from the Connecting the Dots project.
- loc.music.eadmus.mu002008.xml: EAD3 file which contains the source of the outline in the above file. Note that it was a “list” in ead3:bioghist, and the

name was simply changed to “outine” in EAC. I suspect that’s the case for all 5 of the <outline>s from SNAC.

The lone <outline> from Connecting the Dots is an attempt to encode a family tree, which is slightly different from a list, but given that it is just a list, I’d say that the list works fine. That said, being able to encode a family tree would be nice! People also attempt to encode family trees in EAD bioghists (e.g. http://ead-pdfs.library.yale.edu/1674.pdf, which has two family trees “encoded” :))